



# Sufficient conditions for the filtration of a stationary processes to be standard

Gaël Ceillier, Christophe Leuridan

## ► To cite this version:

Gaël Ceillier, Christophe Leuridan. Sufficient conditions for the filtration of a stationary processes to be standard. *Probability Theory and Related Fields*, 2017, 167 (3-4), pp.979-999. 10.1007/s00440-016-0696-2 . hal-01264055

**HAL Id: hal-01264055**

**<https://hal.science/hal-01264055>**

Submitted on 28 Jan 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sufficient conditions for the filtration of a stationary processes to be standard

Gaël Ceillier and Christophe Leuridan

January 28, 2016

## Abstract

Let  $X$  be a stationary process with values in some  $\sigma$ -finite measured state space  $(E, \mathcal{E}, \pi)$ , indexed by  $\mathbb{Z}$ . Call  $\mathcal{F}^X$  its natural filtration. In [3], sufficient conditions were given for  $\mathcal{F}^X$  to be standard when  $E$  is finite, and the proof used a coupling of all probabilities on the finite set  $E$ .

In this paper, we construct a coupling of all laws having a density with regard to  $\pi$ , which is much more involved. Then, we provide sufficient conditions for  $\mathcal{F}^X$  to be standard, generalizing those in [3].

*Mathematics Subject Classification* : 60G10.

*Keywords* : global couplings, filtrations of stationary processes, standard filtrations, generating parametrizations, influence of the distant past.

## 1 Introduction

### 1.1 Setting

In this paper we focus on the natural filtration of stationary processes indexed by the integer line  $\mathbb{Z}$  with values in any  $\sigma$ -finite measured space. By definition, the natural filtration of a process  $X = (X_n)_{n \in \mathbb{Z}}$  is the nondecreasing sequence  $\mathcal{F}^X = (\mathcal{F}_n^X)_{n \in \mathbb{Z}}$  defined by  $\mathcal{F}_n^X = \sigma(X_n^{\triangleleft})$ , where  $X_n^{\triangleleft} = (X_k)_{k \leq n}$  denotes the  $\sigma$ -field generated by the “past of  $X$  at time  $n$ ”. Furthermore, we set  $\mathcal{F}_\infty^X = \sigma(X_k; k \in \mathbb{Z})$  and  $\mathcal{F}_{-\infty}^X = \bigcap_{k \in \mathbb{Z}} \mathcal{F}_k^X$ . All the  $\sigma$ -fields that we consider here are assumed to be complete.

Let  $X = (X_n)_{n \in \mathbb{Z}}$  be any process defined on some probability space. Under loose separability conditions, and up to an enlargement of the probability space, one may assume that  $X$  fulfills a recursion as follows: for every  $n \in \mathbb{Z}$ ,  $X_{n+1}$  is a

function of  $X_n^\triangleleft$  (the “past” of  $X$  at time  $n$ ) and of a “fresh” random variable  $U_{n+1}$ , which brings in some “new” randomness. When this property holds, namely if, for every  $n \in \mathbb{Z}$ , (i)  $U_{n+1}$  is independent of  $\mathcal{F}_n^{X,U}$ , and (ii)  $X_{n+1}$  is measurable with respect to  $\sigma(U_{n+1}) \vee \mathcal{F}_n^X$ , we say that the process  $U = (U_n)_{n \in \mathbb{Z}}$  is a *parametrization* of  $X$  (or of  $\mathcal{F}^X$ ). In particular, the process  $U = (U_n)_{n \in \mathbb{Z}}$  must be independent. Although Schachermayer’s definition [9] of parametrizations imposes that the  $U_n$  are uniform on  $[0, 1]$ , we do not keep that restriction.

Any process indexed by the integer line  $\mathbb{Z}$  with values in some essentially separable space possesses a parametrization (see a proof in [2], section 5.0.5).

Likewise, we say that a parametrization  $U$  is a *generating parametrization* of  $X$ , or that  $U$  *generates*  $X$  if, for every  $n \in \mathbb{Z}$ ,  $X_n$  is measurable with respect to  $\mathcal{F}_n^U$ . This is equivalent to the condition that  $\mathcal{F}_n^X \subset \mathcal{F}_n^U$  for every  $n \in \mathbb{Z}$ , a property which, from now on, we write as  $\mathcal{F}^X \subset \mathcal{F}^U$ .

As a matter of fact, generating parametrizations provide a stronger property which is *immersion* of filtrations. Recall that the filtration  $\mathcal{F}^X$  is *immersed* in the filtration  $\mathcal{F}^U$  if  $\mathcal{F}^X \subset \mathcal{F}^U$  and if, for every  $n \in \mathbb{Z}$ ,  $\mathcal{F}_{n+1}^X$  and  $\mathcal{F}_n^U$  are independent conditionally on  $\mathcal{F}_n^X$ . Roughly speaking, this means that  $\mathcal{F}_n^U$  gives no further information on  $X_{n+1}$  than  $\mathcal{F}_n^X$  does. Equivalently,  $\mathcal{F}^X$  is immersed in  $\mathcal{F}^U$  if every  $\mathcal{F}^X$ -martingale is an  $\mathcal{F}^U$ -martingale. The following easy fact holds (see a proof in [2]).

**Lemma 1.1** *If  $U$  is a generating parametrization of  $X$ , then  $\mathcal{F}^X$  is immersed in  $\mathcal{F}^U$ .*

Another notable property of filtrations is *standardness*. Recall that  $\mathcal{F}^X$  is *standard* if, modulo an enlargement of the probability space, one can immerse  $\mathcal{F}^X$  in a filtration generated by an i.i.d. process with values in a separable space. Vershik introduced standardness in the context of ergodic theory. Examples of non-standard filtrations include the filtrations of  $[T, T^{-1}]$  transformations, introduced in [7]. Split-word processes, inspired by Vershik’s  $(r_n)$ -adic sequences of decreasing partitions [11] and studied in [10] and [8], for instance, also provide non-standard filtrations.

Obviously, lemma 1.1 above implies that if  $X$  has a generating parametrization with values in some essentially separable space, then  $\mathcal{F}^X$  is standard. Whether the converse holds is not known.

Necessary and sufficient conditions for standardness include Vershik’s self-joining criterion and Tsirelson’s notion of  $I$ -cosiness. Both notions are discussed in [6] and are based on conditions which are subtle and not easy to use nor to check in specific cases.

Our goal in this paper is to provide sufficient conditions for the existence of a generating parametrization of  $\mathcal{F}^X$  (and therefore of standardness) that are

easier to use than the ones mentioned above. Each of our conditions involves a quantification of the influence of the distant past of the process on its present. We introduce them in the next section.

## 1.2 Measuring the influence of the distant past

From now on,  $X$  denotes a stationary process indexed by the integer line  $\mathbb{Z}$  with values in some measurable space  $(E, \mathcal{E})$ . We fix a reference measure  $\pi$  on  $E$  which is  $\sigma$ -finite. To introduce some quantities measuring the influence of the past of a process on its present, we need some notations.

**Notation 1** (1) *Slabs:* For any sequence  $(\xi_n)_{n \in \mathbb{Z}}$  in  $E^{\mathbb{Z}}$ , deterministic or random, and any integers  $i \leq j$ ,  $\xi_{i:j}$  is the  $(j - i + 1)$ -uple  $(\xi_n)_{i \leq n \leq j}$  in  $E^{j-i+1}$ .

(2) *Shifts:* If  $k - i = \ell - j$ ,  $\xi_{i:k} = \zeta_{j:\ell}$  means that  $\xi_{i+n} = \zeta_{j+n}$  for every integer  $n$  such that  $0 \leq n \leq k - i$ .

*Infinite case:* Let  $E^{\triangleleft}$  denote the space of sequences  $(\xi_n)_{n \leq -1}$ . For every  $i$  in  $\mathbb{Z}$ , a sequence  $(\xi_n)_{n \leq i}$  is also considered as an element of  $E^{\triangleleft}$  since, similarly to the finite case, one identifies  $\xi_i^{\triangleleft} = (\xi_n)_{n \leq i}$  and  $\zeta_j^{\triangleleft} = (\zeta_n)_{n \leq j}$  if  $\xi_{i+n} = \zeta_{j+n}$  for every integer  $n \leq 0$ .

(3) *Concatenation:* for all  $i \geq 0$ ,  $j \geq 0$ ,  $x = (x_n)_{1 \leq n \leq i}$  in  $E^i$  and  $y = (y_n)_{1 \leq n \leq j}$  in  $E^j$ ,  $xy$  denotes the concatenation of  $x$  and  $y$ , defined as

$$xy = (x_1, \dots, x_i, y_1, \dots, y_j), \quad xy \in E^{i+j}.$$

*Infinite case:* for all  $j \geq 0$ ,  $y = (y_n)_{1 \leq n \leq j}$  in  $E^j$  and  $x = (x_n)_{n \leq -1}$  in  $E^{\triangleleft}$ ,  $xy$  denotes the concatenation of  $x$  and  $y$ , defined as

$$xy = (\dots, x_{-2}, x_{-1}, y_1, \dots, y_j), \quad xy \in E^{\triangleleft}.$$

**Assumption 1** From now on, assume that one can choose a regular version of the conditional law of  $X_0$  given  $X_{-1}^{\triangleleft}$  in such a way that for every  $x \in E^{\triangleleft}$ , the law  $\mathcal{L}(X_0 | X_{-1}^{\triangleleft} = x)$  has a density  $f(\cdot | x)$  with respect to the reference  $\sigma$ -finite measure  $\pi$  on  $E$ . Then, for every  $n \geq 0$  and for every  $x_{1:n} \in E^n$ , the law  $\mathcal{L}(X_0 | X_{-n:-1} = x_{1:n})$  has a density  $f(\cdot | x_{1:n})$  with respect to  $\pi$ .

When  $E$  is countable and  $\pi$  is the counting measure on  $E$ , the densities  $f(\cdot | x)$  coincide with the functions  $p(\cdot | x)$  used in [3] and defined by

$$p(a | x) = \mathbb{P}(X_0 = a \mid X_{-n:-1} = x), \text{ for } x \in E^n \text{ and } a \in E,$$

$$\text{and} \quad p(a | x) = \mathbb{P}(X_0 = a \mid X_{-1}^{\triangleleft} = x), \text{ for } x \in E^{\triangleleft} \text{ and } a \in E.$$

We now introduce three quantities  $\gamma_n$ ,  $\alpha_n$  and  $\delta_n$  measuring the pointwise influence at distance  $n$ .

**Definition 1** For every  $n \geq 0$ , let

$$\gamma_n = 1 - \inf \left\{ \frac{f(a|xz)}{f(a|yz)} ; a \in E, x \in E^\triangleleft, y \in E^\triangleleft, z \in E^n, f(a|yz) > 0 \right\},$$

$$\alpha_n = 1 - \inf_{z \in E^n} \int_E \inf \{f(a|yz); y \in E^\triangleleft\} d\pi(a),$$

$$\delta_n = \sup \{ \|f(\cdot|xz) - f(\cdot|yz)\| ; x \in E^\triangleleft, y \in E^\triangleleft, z \in E^n \},$$

where, for every densities  $f$  and  $g$  on  $E$ ,

$$\|f - g\| = \frac{1}{2} \int_E |f(x) - g(x)| d\pi(x) = \int_E [f(x) - g(x)]_+ d\pi(x).$$

Note that the values  $\gamma_n$ ,  $\alpha_n$  and  $\delta_n$  depend on the choice of the regular version of the conditional law of  $X_0$  given  $X_{-1}^\triangleleft$ . Applying the theorems below requires small influences, so one has to work with a “good” version of the conditional law. Yet, replacing  $\pi$  by an equivalent measure do not modify the quantities  $\gamma_n$ ,  $\alpha_n$  and  $\delta_n$ .

The sequences  $(\gamma_n)_{n \geq 0}$ ,  $(\alpha_n)_{n \geq 0}$  and  $(\delta_n)_{n \geq 0}$  are non-increasing,  $[0, 1]$ -valued, and  $\delta_n \leq \min(\gamma_n, \alpha_n)$  for every  $n \geq 0$  (see the proof in [3], section 5).

For every  $[0, 1]$ -valued sequence  $(\varepsilon_n)_{n \geq 0}$ , we consider the condition

$$\sum_{k=0}^{+\infty} \prod_{n=0}^k (1 - \varepsilon_n) = +\infty. \quad (\mathcal{H}(\varepsilon))$$

For instance,  $\mathcal{H}(\gamma)$  and  $\mathcal{H}(2\delta/(1 + \delta))$  are respectively

$$\sum_{k=0}^{+\infty} \prod_{n=0}^k (1 - \gamma_n) = +\infty, \quad \text{and} \quad \sum_{k=0}^{+\infty} \prod_{n=0}^k \frac{1 - \delta_n}{1 + \delta_n} = +\infty.$$

Observe that if two  $[0, 1]$ -valued sequences  $(\varepsilon_n)_{n \geq 0}$  and  $(\zeta_n)_{n \geq 0}$  are such that  $\varepsilon_n \leq \zeta_n$  for every  $n \geq 0$ , then  $\mathcal{H}(\zeta)$  implies  $\mathcal{H}(\varepsilon)$ . Hence condition  $\mathcal{H}(\varepsilon)$  asserts that  $(\varepsilon_n)_{n \geq 0}$  is small in a certain way.

### 1.3 Statement of the main results

The definition of  $(\gamma_n)_{n \geq 0}$  and the assumption  $\mathcal{H}(\gamma)$  are both in [1]. The main result of [1] is that when  $E$  is of size 2, then  $\mathcal{H}(\gamma)$  implies that  $\mathcal{F}^X$  admits a generating parametrization. This result is restricted by the following three conditions. First, the size of  $E$  must be 2. Second, one must control the *ratios*

of probabilities which define  $\gamma_n$ . Third,  $\mathcal{H}(\gamma)$  implies that  $\gamma_0 < 1$ , therefore one can show that  $\mathcal{H}(\gamma)$  implies the existence of  $c > 0$  such that  $f(a|x) \geq c$  for every  $x$  in  $E^\triangleleft$  and  $a$  in  $E$ .

Theorem 2 in [3] improves on this and gets rid of the first two restrictions. And Theorem 1 below improves on Theorem 2 in [3].

**Theorem 1** (1) *If  $E$  is finite and if  $\mathcal{H}(2\delta/(1+\delta))$  holds, then  $\mathcal{F}^X$  admits a generating parametrization.*

(2) *If the size of  $E$  is 2 and if  $\mathcal{H}(\delta)$  holds, then  $\mathcal{F}^X$  admits a generating parametrization.*

Theorem 1 brings two improvements on the first result in [3]: the assumption  $\mathcal{H}(2\delta)$  is replaced by  $\mathcal{H}(2\delta/(1+\delta))$  and the extra hypothesis  $\delta_0 < 1/2$  is not required anymore.

Another measure of influence, based on the quantities  $\alpha_n$ , is introduced and used in [5] (actually the notation there is  $a_n = 1 - \alpha_n$ ). The authors show that if  $\mathcal{H}(\alpha)$  holds, there exists a perfect sampling algorithm for the process  $X$ , a result which implies that  $\mathcal{F}^X$  admits a generating parametrization.

Our Theorem 1, the result in [1] and the exact sampling algorithm of [5] all require an upper bound of some pointwise influence sequence. The second theorem in [3] uses a less restrictive hypothesis based on some average influences  $\eta_n$ , defined below.

**Definition 2** *For every  $n \geq 0$ , let  $\eta_n$  denote the average influence at distance  $n$ , defined as*

$$\eta_n = \mathbb{E}[\|f(\cdot|Y_{-n:-1}) - f(\cdot|X_{-n-1}^\triangleleft Y_{-n:-1})\|],$$

*where  $Y$  is an independent copy of  $X$ , and call  $\mathcal{H}'(\eta)$  the condition*

$$\sum_{k=0}^{+\infty} \eta_k < +\infty. \quad (\mathcal{H}'(\eta))$$

As before, the definition of  $\eta_n$  depends on the choice of the regular version of the conditional law of  $X_0$  given  $X_{-1}^\triangleleft$ , but it does not depend on the choice of the reference measure  $\pi$ . We recall the result from [3].

**Theorem 2 (Theorem 3 of [3])** *Assume that  $E$  is finite and that for every  $a$  in  $E$ ,  $p(a|X_{-1}^\triangleleft) > 0$  almost surely (priming condition). Then,  $\mathcal{H}'(\eta)$  implies that  $\mathcal{F}^X$  admits a generating parametrization.*

Note that the sequence  $(\eta_n)_{n \geq 0}$  is  $[0, 1]$ -valued. If  $\eta_n < 1$  for every  $n \leq 0$ , then  $\mathcal{H}'(\eta)$  implies  $\mathcal{H}(\eta)$ . Yet, since  $\eta_n \leq \delta_n$  for every  $n \geq 0$ , the condition  $\mathcal{H}'(\eta)$  cannot be compared to the conditions  $\mathcal{H}(\delta)$  and  $\mathcal{H}(2\delta/(1+\delta))$ .

The main result of the present paper is the theorem below which extends theorem 3 of [3] to any  $\sigma$ -finite measured space  $(E, \mathcal{E}, \pi)$ .

**Theorem 3** *Let  $(E, \mathcal{E}, \pi)$  be any  $\sigma$ -finite measured space. Assume that one has chosen a regular version of the conditional law of  $X_0$  given  $X_{-1}^\triangleleft$  in such a way that for every  $x \in E^\triangleleft$ , the law  $\mathcal{L}(X_0|X_{-1}^\triangleleft = x)$  has a density  $f(\cdot|x)$  with respect to  $\pi$  and that for  $\pi$ -almost every  $a$  in  $E$ ,  $f(a|X_{-1}^\triangleleft) > 0$  almost surely (priming condition). Then,  $\mathcal{H}'(\eta)$  implies that  $\mathcal{F}^X$  admits a generating parametrization, hence is standard.*

In this paper, we call *priming condition* the condition “for  $\pi$ -almost every  $a$  in  $E$ ,  $f(a|X_{-1}^\triangleleft) > 0$  almost surely”. This generalises the priming condition of [3]. Indeed, when  $E$  is finite and  $\pi$  is the counting measure on  $E$ , the priming condition is equivalent to the assumption that for every  $a$  in  $E$ ,  $p(a|X_{-1}^\triangleleft) > 0$  almost surely, that is, the condition involved in theorem 2. An equivalent form, which is actually the one used in the proof of theorem 2, is

$$\inf_{a \in E} p(a|X_{-1}^\triangleleft) > 0 \text{ almost surely.}$$

But this condition cannot be satisfied anymore when  $E$  is infinite. Therefore, the proof of theorem 3 requires new ideas to use the priming condition, see our priming lemma in section 3 (lemma 3.1).

## 1.4 Global couplings

A key tool for the construction of our parametrization is a new global coupling. This global coupling, which applies when  $E$  is infinite, is much more involved than the coupling introduced in [3] when  $E$  is finite.

Recall that if  $p$  and  $q$  are two *fixed* probabilities on a countable space  $E$ , then for every random variables  $Z_p$  and  $Z_q$  with laws  $p$  and  $q$  defined on the same probability space,

$$\mathbb{P}[Z_p \neq Z_q] \geq \|p - q\|,$$

where  $\|p - q\|$  is the distance in total variation between  $p$  and  $q$ , defined as

$$\|p - q\| = \frac{1}{2} \sum_{a \in E} |p(a) - q(a)| = \sum_{a \in E} [p(a) - q(a)]_+.$$

Conversely, a standard construction in coupling theory provides some random variables  $Z_p$  and  $Z_q$  with laws  $p$  and  $q$  such that  $\mathbb{P}[Z_p \neq Z_q] = \|p - q\|$ .

Couplings can be generalized to any set of laws, on a measurable space  $(E, \mathcal{E})$ .

**Definition 3** Note  $\mathcal{P}(E)$  the set of probabilities on the measurable space  $(E, \mathcal{E})$ . Let  $\mathcal{A}$  be a subset of  $\mathcal{P}(E)$ . A global coupling for  $\mathcal{A}$  is a random variable  $U$  with values in some measurable space  $(F, \mathcal{F})$  and a function  $g$  from  $F \times \mathcal{A}$  to  $E$  such that for every  $p \in \mathcal{A}$ ,  $g(U, p)$  is a random variable of law  $p$  on  $E$ .

Good global couplings are such that for every  $p$  and  $q$  in  $\mathcal{A}$ , the probability  $\mathbb{P}[g(U, p) \neq g(U, q)]$  is small when  $\|p - q\|$  is small.

When  $E = \{0, 1\}$ , a classical coupling used in [1] is given by  $g(U, p) = \mathbf{1}_{\{U \leq p(1)\}}$ , where  $U$  is a uniform random variable on  $[0, 1]$ . This coupling satisfies the equality  $\mathbb{P}[g(U, p) \neq g(U, q)] = \|p - q\|$ , for every  $p$  and  $q$  in  $\mathcal{P}(E)$ .

This coupling has been extended to any finite set in [3]. Furthermore, the construction can still be extended to any countable set, as follows.

**Proposition 1.2** Assume that  $E$  is countable. Let  $\varepsilon = (\varepsilon_a)_{a \in E}$  be an i.i.d. family of exponential random variables with parameter 1. Then, for any probability  $p$  on  $E$ , there exists almost surely one and only one  $a \in E$  such that

$$\frac{\varepsilon_a}{p(a)} = \inf_{b \in E} \frac{\varepsilon_b}{p(b)} \quad (\star),$$

with the convention that  $\varepsilon_b/p(b) = +\infty$  if  $p(b) = 0$ .

- Define  $g(\varepsilon, p)$  (almost surely) as the only index  $a$  such that  $(\star)$  holds. Then the law of the random variable  $g(\varepsilon, p)$  is  $p$ .
- For every  $p, q$  in  $\mathcal{P}(E)$ ,

$$\mathbb{P}[g(\varepsilon, p) \neq g(\varepsilon, q)] \leq 2 \frac{\|p - q\|}{1 + \|p - q\|} \leq 2 \|p - q\|.$$

The bound  $2\|p - q\|/(1 + \|p - q\|)$  improves on the bound  $2\|p - q\|$  given in proposition 2.6 of [3] and explains that the assumption  $\mathcal{H}(2\delta/(1+\delta))$  of theorem 1 is weaker than the assumptions  $\mathcal{H}(2\delta)$  and  $\delta_0 < 1/2$  required in theorem 2 of [3], although the same proof works. The proof of proposition 1.2 will be given in section 5.

We now briefly introduce a new global coupling that works on any measured  $\sigma$ -finite space  $(E, \mathcal{E}, \pi)$ . The construction and the properties of the coupling  $(g, U)$  will be detailed in section 2.

Let  $\mathcal{P}_\pi(E)$  be the subset of  $\mathcal{P}(E)$  of all probability laws  $p$  on  $E$  admitting a density  $f_p$  with respect to  $\pi$ . Let  $\lambda$  be the Lebesgue measure on  $\mathbb{R}^+$ . Let  $U$  be a Poisson point process on  $E \times \mathbb{R}^+ \times \mathbb{R}^+$  of intensity  $\pi \otimes \lambda \otimes \lambda$ .

In the following, any element  $(x, y, t)$  of  $U$  is seen as a point  $(x, y)$  in  $E \times \mathbb{R}^+$  appearing at the time  $t$ . Since  $\lambda$  is diffuse, the third components of the elements



of  $U$  are almost surely all distinct. For any density function  $f$  on  $(E, \mathcal{E}, \pi)$ , set  $D_f = \{(x, y) \in E \times \mathbb{R}^+ : y \leq f(x)\}$  and call  $t_f$  the instant of appearance of a point under the graph of  $f$ . Then, almost surely,  $t_f$  is positive and finite and the set  $U \cap (D_f \times [0, t_f])$  is almost surely reduced to a single point  $(x_f, y_f, t_f)$ . The random variable  $x_f$  has density  $f$  with respect to  $\pi$ .

**Proposition 1.3** *With the notations above, the formula  $g(U, p) = x_{f_p}$  defines a global coupling for  $\mathcal{P}_\pi(E)$  such that for every  $p$  and  $q$  in  $\mathcal{P}_\pi(E)$ ,*

$$\mathbb{P}[g(U, p) \neq g(U, q)] \leq \frac{2\|p - q\|}{1 + \|p - q\|} \leq 2\|p - q\|.$$

The paper is organised as follows : In section 2, we detail the construction of the global coupling just mentionned, we prove proposition 1.3 and we construct a parametrization of the process  $(X_n)_{n \in \mathbb{Z}}$ . Section 3 is devoted to the proof of a key intermediate result which we call priming lemma, which uses the priming condition. In section 4, we complete the proof of theorem 3, showing that the parametrization  $(U_n)_{n \in \mathbb{Z}}$  constructed in section 2 generates the filtration  $\mathcal{F}^X$ .

## 2 Construction of a parametrization with a global coupling

### 2.1 Construction of a global coupling

Let  $(E, \mathcal{E}, \pi)$  be a  $\sigma$ -finite measured space. Let  $\lambda$  be the Lebesgue measure on  $\mathbb{R}^+$ . Set  $\mu = \pi \otimes \lambda$  on  $E \times \mathbb{R}^+$ . Denote by  $\mathcal{D}$  the set of countable subsets of  $E \times \mathbb{R}^+ \times \mathbb{R}^+$ .

Let  $U$  be a Poisson point process on  $E \times \mathbb{R}^+ \times \mathbb{R}^+$  of intensity  $\pi \otimes \lambda \otimes \lambda$ . By definition,  $U$  is a  $\mathcal{D}$ -valued random variable such that for any pairwise disjoint sets  $B_k$  in  $\mathcal{E}$ , of finite measure for  $\mu \otimes \lambda = \pi \otimes \lambda \otimes \lambda$ , the cardinalities  $|U \cap B_k|$  are independent Poisson random variables with respective parameters  $(\mu \otimes \lambda)(B_k)$ .

In the following, any element  $(x, y, t)$  of  $U$  is seen as a point  $(x, y)$  in  $E \times \mathbb{R}^+$  appearing at time  $t$ . Since  $\lambda$  is diffuse, the third components of the elements of  $U$  are almost surely all distinct.

**Notation 2** *For any measurable subset  $A$  of  $E \times \mathbb{R}^+$ , the time of appearance of a point in  $A$  is*

$$t_A(U) = \inf \{t \geq 0 : U \cap (A \times [0, t]) \neq \emptyset\}.$$

Let us recall some classical and useful properties of Poisson point processes. If the measure  $\mu(A)$  is positive and finite, the random variable  $t_A(U)$  has exponential law of parameter  $\mu(A)$ . Moreover the set  $U \cap (A \times [0, t_A(U)])$  is almost surely reduced to a single point  $(x_A(U), y_A(U), t_A(U))$ . The couple  $(x_A(U), y_A(U))$  thus defined is a random variable of law  $\mu(\cdot|A) = \mu(\cdot \cap A)/\mu(A)$  which is independent from  $t_A(U)$ .

Moreover if the sets  $B_k$  are pairwise disjoint measurable subsets of  $E \times \mathbb{R}^+$ , the random variables  $(x_{B_k}(U), y_{B_k}(U), t_{B_k}(U))$  are independent.

The coordinates  $x_A(U), y_A(U), t_A(U)$  vary as a function of the variable  $U$ . This notation will be used later on when we will consider several Poisson point processes, but, by abuse of notation it will be abbreviated to  $x_A, y_A, t_A$  when there is no ambiguity on the Poisson point process.

**Lemma 2.1** *If the  $n$  sets  $A_k$  are measurable subsets of  $E \times \mathbb{R}^+$ , of finite positive measures for  $\mu$ , then almost surely*

$$t_{A_1} = \dots = t_{A_n} \iff t_{A_1 \cap \dots \cap A_n} = t_{A_1 \cup \dots \cup A_n}.$$

*Proof* The reverse implication follows from the inequalities

$$t_{A_1 \cup \dots \cup A_n} = \min\{t_{A_1}, \dots, t_{A_n}\} \leq \max\{t_{A_1}, \dots, t_{A_n}\} \leq t_{A_1 \cap \dots \cap A_n}.$$

The direct implication follows from the fact that, almost surely, the third components of the elements of  $U$  are all distinct. Therefore if  $t_{A_1} = \dots = t_{A_n}$ , then the point  $(x_{A_k}, y_{A_k}, t_{A_k})$  does not depend on  $k$ , therefore  $(x_{A_k}, y_{A_k})$  belongs to  $A_1 \cap \dots \cap A_n$ . This ends the proof.  $\square$

We now study the dependance of  $x_A$  with respect to  $A$ .

**Proposition 2.2** *Let  $A$  and  $B$  be two measurable subsets of  $E \times \mathbb{R}^+$  of measures finite and positive, one has that*

$$\{x_A = x_B\} \supset \{t_A = t_B\} \text{ almost surely}$$

and

$$\mathbb{P}[t_A = t_B] = \frac{\mu(A \cap B)}{\mu(A \cup B)}.$$

*Proof* Since the third components of the elements of  $U$  are almost surely all distinct, one has

$$t_A = t_B \Rightarrow x_A = x_B \text{ almost surely.}$$

By lemma 2.1, and the equality  $t_{A \cup B} = \min\{t_{A \cap B}, t_{A \Delta B}\}$ ,

$$\begin{aligned} t_A = t_B &\iff t_{A \cap B} = t_{A \cup B} \\ &\iff t_{A \cap B} \leq t_{A \Delta B}. \end{aligned}$$

Since  $t_{A \cap B}$  and  $t_{A \Delta B}$  are independent exponential random variables with parameters  $\mu(A \cap B)$  and  $\mu(A \Delta B)$ , one gets

$$\mathbb{P}[t_A = t_B] = \mu(A \cap B) / \mu(A \cup B),$$

which completes the proof.  $\square$

**Remark 1** *Note that if  $\pi$  is diffuse, then  $\{x_A = x_B\} = \{t_A = t_B\}$  almost surely, thus  $\mathbb{P}[x_A = x_B] = \mu(A \cap B) / \mu(A \cup B)$ .*

Let us introduce some subsets to which we will apply proposition 2.2.

**Notation 3** *For any measurable map  $f : E \rightarrow \mathbb{R}^+$ , denote by  $D_f$  the part of  $E \times \mathbb{R}^+$  located below the graph of  $f$ :*

$$D_f = \{(x, y) \in E \times \mathbb{R}^+ : y \leq f(x)\}.$$

Then,

$$\mu(D_f) = \int_E f(x) \, d\pi(x).$$

If this measure is positive and finite, then we denote in abbreviated form :

$$(x_f, y_f, t_f) = (x_{D_f}, y_{D_f}, t_{D_f}).$$

Proposition 1.3 is a direct consequence of the next lemma.

**Lemma 2.3** *Let  $\alpha > 0$  and  $f$  be a probability density function on  $(E, \pi)$ . Then*

- *the random variable  $x_{\alpha f}$  has density  $f$  with respect to  $\pi$ . Thus  $(U, x)$  is a global coupling for the set of all laws on  $E$  with a density with respect to  $\pi$*
- *for all probabilities  $f$  and  $g$  on  $E$ ,*

$$\mathbb{P}[x_f = x_g] \geq \mathbb{P}[t_f = t_g] = \frac{1 - \|f - g\|}{1 + \|f - g\|}.$$

*Proof* The first result follows from the fact that the law of  $(x_{\alpha f}, y_{\alpha f})$  is  $\mu(\cdot | D_f)$ . The second result follows from proposition 2.2.  $\square$

## 2.2 Parametrizing one random variable

Let  $X$  be a random variable with values in  $E$  with a density  $f$  with respect to  $\pi$ . Let  $W$  be a Poisson point process on  $E \times \mathbb{R}^+ \times \mathbb{R}^+$  of intensity  $\pi \otimes \lambda \otimes \lambda$ , and  $V$  be a random variable uniform on  $[0, 1]$ , such that  $X, V, W$  are independent. Let us modify  $W$  in such a way that the first two components of the first point which appears in  $D_f$  are  $(X, Vf(X))$ . We set

$$U = \left[ W \setminus \{(x_f(W), y_f(W), t_f(W))\} \right] \cup \{(X, Vf(X), t_f(W))\}.$$

**Proposition 2.4** *The process  $U$  thus defined is a Poisson point process on  $E \times \mathbb{R}^+ \times \mathbb{R}^+$  of intensity  $\pi \otimes \lambda \otimes \lambda$ , such that  $x_f(U) = X$ .*

*Proof* One has directly  $x_f(U) = X$ . Let us show that  $U$  is a Poisson point process. Denote by  $\overline{D_f}$  the complementary of  $D_f$  in  $E \times \mathbb{R}^+$ . Let us split  $W$  into two independent Poisson point processes  $W_1$  and  $W_2$  by setting :

$$W_1 = W \cap (D_f \times \mathbb{R}^+) \text{ and } W_2 = W \cap (\overline{D_f} \times \mathbb{R}^+).$$

In a similar way, set

$$U_1 = U \cap (D_f \times \mathbb{R}^+) \text{ and } U_2 = U \cap (\overline{D_f} \times \mathbb{R}^+).$$

It suffices to show that  $\mathcal{L}(U_1, U_2) = \mathcal{L}(W_1, W_2)$ . Note that  $U_2 = W_2$  and  $U_1$  is a function of  $X, V, W_1$ , therefore  $U_1$  is independent from  $U_2$  and  $\mathcal{L}(U_2) = \mathcal{L}(W_2)$ . Thus what remains to be proved is that  $\mathcal{L}(U_1) = \mathcal{L}(W_1)$ .

Let  $(w_k, t_k)_{k \geq 1}$  be the sequence of the points of  $W_1$  in  $D_f \times \mathbb{R}^+$  ordered in such a way that  $(t_k)_{k \geq 1}$  is an increasing sequence. Then

$$U_1 = \bigcup_{k \geq 1} \{(u_k, t_k)\}$$

with  $u_1 = (X, Vf(X))$  and  $u_k = w_k$  for  $k \geq 2$ . The random variables  $(w_k)_{k \geq 1}$  form an i.i.d. sequence of law  $\mu(\cdot | D_f)$ , independent of  $((t_k)_{k \geq 1}, X, V)$ .

A direct computation shows that the random variable  $u_1 = (X, Vf(X))$  has law  $\mu(\cdot | D_f)$ . Moreover,  $u_1$  is independent of  $((w_k)_{k \geq 2}, (t_k)_{k \geq 1})$  by independence of  $X, V$  and  $W$ . Therefore  $\mathcal{L}(U_1) = \mathcal{L}(W_1)$ , which completes the proof.  $\square$

## 2.3 Construction of a parametrization

Assume that  $(X_n)_{n \in \mathbb{Z}}$  is a stationary process with values in  $E$  such that given  $X_{-1}^{\leq}, X_0$  admits conditional densities  $(f(\cdot | x))_{x \in E^{\leq}}$  with respect to  $\pi$ . One wants to apply the previous construction at each time  $n \in \mathbb{Z}$ .

Recall that for every  $n \in \mathbb{Z}^+$  and  $x \in E^{\mathbb{Z}_*^-}$  the function  $f(\cdot|x)$  is the density of the law  $\mathcal{L}(X_0|X^\triangleleft = x)$  with respect to  $\pi$ .

Let  $(W_n)_{n \in \mathbb{Z}}$  be a sequence of independent Poisson point processes on  $E \times \mathbb{R}^+ \times \mathbb{R}^+$  of intensity  $\pi \otimes \lambda \otimes \lambda$ , and let  $(V_n)_{n \in \mathbb{Z}}$  be a sequence of independent and uniform random variables on  $[0, 1]$ , such that the sequences  $(X_n)_{n \in \mathbb{Z}}$ ,  $(V_n)_{n \in \mathbb{Z}}$ ,  $(W_n)_{n \in \mathbb{Z}}$  are independent. Set

$$U_n = W_n \cup \{(X_n, V_n f_{n-1}(X_n), t_{f_{n-1}}(W_n))\} \setminus \{(x_{f_{n-1}}(W_n), y_{f_{n-1}}(W_n), t_{f_{n-1}}(W_n))\},$$

where  $f_{n-1} = f(\cdot|X_{n-1}^\triangleleft)$ . Since  $X$  is stationary,  $f_{n-1}$  is the law of  $X_n$  given the past  $X_{n-1}^\triangleleft$ . Proposition 2.4 implies that for every  $n \in \mathbb{Z}$ ,  $U_n$  is a Poisson point process independent from  $\mathcal{F}_{n-1}^{X,V,W}$ . Moreover,  $X_n = x_{f_{n-1}}(U_n)$  and  $U_n$  is a function of  $W_n$ ,  $V_n$  and  $X_n^\triangleleft$ . Therefore we have the following result.

**Proposition 2.5** *The process  $(U_n)_{n \in \mathbb{Z}}$  thus defined is an i.i.d. sequence of Poisson point processes on  $E \times \mathbb{R}^+ \times \mathbb{R}^+$  of intensity  $\pi \otimes \lambda \otimes \lambda$ , which is a parametrization of  $\mathcal{F}^X$  with the recursion formula  $X_n = x_{f_{n-1}}(U_n)$ .*

### 3 Priming lemma

Fix an integer  $\ell \geq 0$ . The following lemma provides a random variable  $Z_{1:\ell}$  and an event  $H_\ell$  which both depend only on  $U_{1:\ell} = (U_1, \dots, U_\ell)$  such that when  $H_\ell$  occurs,  $X_{1:\ell} = Z_{1:\ell}$  with probability close to 1.

**Lemma 3.1 (Priming lemma)** *Assume that for  $\pi$ -almost every  $a$  in  $E$ , we have  $f(a|X_{-1}^\triangleleft) > 0$  almost surely. For every  $\varepsilon > 0$  and  $\ell \geq 0$ , there exists a random variable  $Z_{1:\ell} = (Z_1, \dots, Z_\ell)$  with values in  $E^\ell$  and an event  $H_\ell$ , which are both functions of  $U_{1:\ell} = (U_1, \dots, U_\ell)$  only, such that*

- $\mathbb{P}[H_\ell] > 0$ ,
- $\mathcal{L}(Z_{1:\ell}) = \mathcal{L}(X_{1:\ell})$ ,
- $H_\ell$  is independent of  $Z_{1:\ell}$ ,
- $\mathbb{P}[X_{1:\ell} \neq Z_{1:\ell} \mid H_\ell] \leq \varepsilon$ .

Note that the hypothesis and the conclusions of lemma 3.1 are not modified if one replace  $\pi$  by an equivalent measure. Thus in the proof, one may assume without loss of generality that  $\pi$  is a probability measure.

The proof proceeds by induction. For  $\ell = 0$  the result is trivial with  $H_0 = \Omega$  and  $Z_0$  equal to the empty word. The inductive step from  $\ell$  to  $\ell + 1$  uses the next two lemmas. These lemmas show that with a probability close to 1, most of the graph of  $f(\cdot|X_\ell^\triangleleft)$  is between  $m^{-1}f(\cdot|Z_{1:\ell})$  and  $n$  for suitable constants  $m$  and  $n$ .

**Lemma 3.2** *Let  $\ell \geq 0$  be an integer,  $H$  an event of positive probability, and  $Z$  be a random variable with values in  $E^\ell$ . For every  $\varepsilon > 0$ , there exists a real number  $m \geq 1$  such that*

$$\mathbb{E} \left[ \int_E [f(a|Z) - mf(a|X_\ell^\triangleleft)]_+ d\pi(a) \mid H \right] \leq \varepsilon.$$

*Proof.* The priming condition and the stationarity of  $X$  ensure that, for  $\pi$ -almost every  $a \in E$ ,  $f(a|X_\ell^\triangleleft) > 0$  almost surely, so

$$[f(a|Z) - mf(a|X_\ell^\triangleleft)]_+ \rightarrow 0 \text{ a.s. as } m \rightarrow +\infty.$$

Hence, by dominated convergence

$$\mathbb{E} \left[ \int_E [f(a|Z) - mf(a|X_\ell^\triangleleft)]_+ d\pi(a) \right] \rightarrow 0 \text{ as } m \rightarrow +\infty.$$

The same result holds with  $\mathbb{E}[\cdot|H]$  instead of  $\mathbb{E}$  since  $\mathbb{P}[\cdot|H] \leq \mathbb{P}[H]^{-1}\mathbb{P}[\cdot]$ , which concludes the proof.  $\square$

**Lemma 3.3** *Let  $\ell \geq 0$  be an integer,  $m \geq 1$  a real number,  $H$  an event of positive probability and  $Z$  be a random variable with values in  $E^\ell$ . For every  $\varepsilon > 0$ , there exists a real number  $n \geq 1$  such that*

$$\mathbb{E} \left[ \int_E [f(a|X_\ell^\triangleleft) - n]_+ d\pi(a) \mid H \right] \leq \varepsilon.$$

*For such a real number  $n$ , there exists a random variable  $M$ , with values in  $[n, n+1]$ , which is a function of  $Z$  only, such that*

$$\int_E \sup(m^{-1}f(x|Z), M) d\pi(x) = n+1.$$

*Since  $M \geq n$ , one has  $[f(a|X_\ell^\triangleleft) - M]_+ \leq [f(a|X_\ell^\triangleleft) - n]_+$ , so*

$$\mathbb{E} \left[ \int_E [f(a|X_\ell^\triangleleft) - M]_+ d\pi(a) \mid H \right] \leq \varepsilon.$$

*Proof.* The same method as in the proof of lemma 3.2 provides a real number  $n \geq 1$  such that

$$\mathbb{E} \left[ \int_E [f(a|X_\ell^\triangleleft) - n]_+ d\pi(a) \mid H \right] \leq \varepsilon.$$

Define a random application  $\phi$  from  $\mathbb{R}^+$  to  $\mathbb{R}^+$  by

$$\phi(s) = \int_E \sup(m^{-1}f(x|X_\ell^\triangleleft), s) d\pi(x).$$

Since for every  $s \in \mathbb{R}^+$ ,  $s \leq \sup(m^{-1}f(x|X_\ell^\triangleleft), s) \leq m^{-1}f(x|X_\ell^\triangleleft) + s$ , one has  $s \leq \phi(s) \leq m^{-1} + s \leq 1 + s$  (recall that  $\pi$  is assumed to be a probability). Hence  $\phi(n) \leq n + 1 \leq \phi(n + 1)$ . Then, since  $\phi$  is continuous, the random variable

$$M = \inf\{s \in \mathbb{R}^+ : \phi(s) = n + 1\}$$

is well defined and satisfies the conclusion of the lemma.  $\square$

We can now prove the induction step of the proof of lemma 3.1.

*Proof.* Let  $\varepsilon > 0$  and  $\ell \in \mathbb{N}$ . Assume that one has constructed  $H_\ell$  and  $Z_{1:\ell}$ , which are both functions of  $U_{1:\ell}$  only, and such that

- $\mathbb{P}[H_\ell] > 0$ ,
- $\mathcal{L}(Z_{1:\ell}) = \mathcal{L}(X_{1:\ell})$ ,
- $H_\ell$  is independent of  $Z_{1:\ell}$ ,
- $\mathbb{P}[Z_{1:\ell} \neq X_{1:\ell} \mid H_\ell] \leq \varepsilon/3$ .

Lemmas 3.2 and 3.3, applied to  $\ell$ ,  $H = H_\ell$ ,  $Z = Z_{1:\ell}$  and  $\varepsilon/3$  provide two real numbers  $m \geq 1$ ,  $n \geq 1$  and a  $\sigma(Z_{1:\ell})$ -measurable random variable  $M$ . Set

$$f_\ell = f(\cdot|X_\ell^\triangleleft), \quad f'_\ell = f(\cdot|Z_{1:\ell}), \quad A = D_{m^{-1}f'_\ell}, \quad B = D_M, \quad C = D_{f_\ell}.$$

The random sets  $A$  and  $B$  are  $\sigma(Z_{1:\ell})$ -measurable, therefore  $\sigma(U_{1:\ell})$ -measurable, whereas  $C$  is  $\mathcal{F}_\ell^X$ -measurable. Moreover, by lemma 3.3,

$$\mu(A \cup B) = \int_E \sup(m^{-1}f(a|Z_{1:\ell}), M) d\pi(a) = n + 1,$$

whereas

$$\mu(A) = \int_E m^{-1}f(a|Z_{1:\ell}) d\pi(a) = m^{-1}.$$

Let

$$H' = \{t_A(U_{\ell+1}) \leq t_B(U_{\ell+1})\} = \{t_A(U_{\ell+1}) \leq t_{B \setminus A}(U_{\ell+1})\},$$

$$H_{\ell+1} = H_\ell \cap H', \quad Z_{\ell+1} = x_A(U_{\ell+1}),$$

and recall that  $X_{\ell+1} = x_{f_\ell}(U_{\ell+1}) = x_C(U_{\ell+1})$  by proposition 2.5.

For every  $F \in \mathcal{E}^{\otimes \ell}$  and  $G \in \mathcal{E}$ ,

$$\mathbb{P}[Z_{1:\ell+1} \in F \times G ; H_{\ell+1} \mid \mathcal{F}_\ell^{X,U}] = \mathbf{1}_F(Z_{1:\ell}) \mathbf{1}_{H_\ell} \mathbb{P}[Z_{\ell+1} \in G ; H' \mid \mathcal{F}_\ell^{X,U}]. \quad (1)$$

Conditionally on  $\mathcal{F}_\ell^{X,U}$ , the random variables  $x_A(U_{\ell+1})$ ,  $t_A(U_{\ell+1})$  and  $t_{B \setminus A}(U_{\ell+1})$  are independent, with respective laws  $f'_\ell \cdot \pi$ , exponential with parameter  $\mu(A)$ , exponential with parameter  $\mu(B \setminus A)$ . Therefore,

$$\mathbb{P}[Z_{\ell+1} \in G ; H' \mid \mathcal{F}_\ell^{X,U}] = \frac{\mu(A)}{\mu(A \cup B)} \int_G f'_\ell d\pi = \frac{1}{m(n+1)} \int_G f'_\ell d\pi. \quad (2)$$

In particular,  $\mathbb{P}[H' \mid \mathcal{F}_\ell^{X,U}] = \frac{1}{m(n+1)}$ , so  $H'$  is independent of  $\mathcal{F}_\ell^{X,U}$ ; this fact will be used at the end of the proof.

Combining equalities 1 and 2, taking expectations and using the induction hypothesis yield

$$\begin{aligned} \mathbb{P}[Z_{1:\ell+1} \in F \times G ; H_{\ell+1}] &= \frac{1}{m(n+1)} \mathbb{E} \left[ \mathbf{1}_F(Z_{1:\ell}) \mathbf{1}_{H_\ell} \int_G f(\cdot \mid Z_{1:\ell}) d\pi \right] \\ &= \frac{\mathbb{P}(H_\ell)}{m(n+1)} \mathbb{E} \left[ \mathbf{1}_F(X_{1:\ell}) \int_G f(\cdot \mid X_{1:\ell}) d\pi \right] \\ &= \frac{\mathbb{P}(H_\ell)}{m(n+1)} \mathbb{P}[X_{1:\ell+1} \in F \times G]. \end{aligned}$$

This shows that  $Z_{1:\ell+1}$  and  $H_{\ell+1}$  are independent, that  $\mathcal{L}(Z_{1:\ell+1}) = \mathcal{L}(X_{1:\ell+1})$  and that  $\mathbb{P}(H_{\ell+1}) = \mathbb{P}(H_\ell)/(m(n+1)) > 0$ .

What remains to be proved is the upper bound of  $\mathbb{P}[Z_{\ell+1} \neq X_{\ell+1} \mid H_{\ell+1}]$ . By proposition 2.2 and lemma 2.1,

$$\begin{aligned} \{X_{\ell+1} = Z_{\ell+1}\} \cap H' &\supset \{t_C(U_{\ell+1}) = t_A(U_{\ell+1}) = t_{A \cup B}(U_{\ell+1})\} \\ &= \{t_{A \cap C \cap (A \cup B)}(U_{\ell+1}) = t_{A \cup C \cup (A \cup B)}(U_{\ell+1})\} \\ &= \{t_{A \cap C}(U_{\ell+1}) = t_{A \cup C \cup B}(U_{\ell+1})\}. \end{aligned}$$

Since  $A$ ,  $B$  and  $C$  are measurable for  $\mathcal{F}_\ell^{X,U}$  and  $U_{\ell+1}$  is independent from  $\mathcal{F}_\ell^{X,U}$ , one has

$$\mathbb{P}[X_{\ell+1} = Z_{\ell+1}; H' \mid \mathcal{F}_\ell^{X,U}] \geq \frac{\mu(A \cap C)}{\mu(A \cup C \cup B)}.$$

Moreover

$$\begin{aligned} \mu(A \cap C) &= \frac{1}{m} \int_E \min(f(a \mid Z_{1:\ell}), mf(a \mid X_\ell^\triangleleft)) d\pi(a) \\ &= \frac{1}{m} \left( \int_E f(a \mid Z_{1:\ell}) d\pi(a) - \int_E [f(a \mid Z_{1:\ell}) - mf(a \mid X_\ell^\triangleleft)]_+ d\pi(a) \right) \\ &= \frac{1}{m} \left( 1 - \int_E [f(a \mid Z_{1:\ell}) - mf(a \mid X_\ell^\triangleleft)]_+ d\pi(a) \right) \end{aligned}$$



and

$$\begin{aligned}
\mu(A \cup C \cup B) &= \int_E \sup(m^{-1}f(a|Z_{1:\ell}), f(a|X_\ell^\triangleleft), M) d\pi(a) \\
&= \int_E \sup(m^{-1}f(a|Z_{1:\ell}), M) d\pi(a) \\
&\quad + \int_E [f(a|X_\ell^\triangleleft) - \sup(m^{-1}f(a|Z_{1:\ell}), M)]_+ d\pi(a) \\
&\leq n+1 + \int_E [f(a|X_\ell^\triangleleft) - M]_+ d\pi(a) \\
&\leq (n+1) \left(1 + \int_E [f(a|X_\ell^\triangleleft) - M]_+ d\pi(a)\right),
\end{aligned}$$

so

$$\mu(A \cup C \cup B)^{-1} \geq \frac{1}{n+1} \left(1 - \int_E [f(a|X_\ell^\triangleleft) - M]_+ d\pi(a)\right).$$

Thus

$$\begin{aligned}
m(n+1)\mathbb{P}[X_{\ell+1} = Z_{\ell+1}; H' \mid \mathcal{F}_\ell^{X,U}] &\geq m(n+1) \frac{\mu(A \cap C)}{\mu(A \cup C \cup B)} \\
&\geq 1 - \int_E [f(a|X_\ell^\triangleleft) - M]_+ d\pi(a) \\
&\quad - \int_E [f(a|Z_{1:\ell}) - mf(a|X_\ell^\triangleleft)]_+ d\pi(a).
\end{aligned}$$

Since  $H_\ell \in \mathcal{F}_\ell^{X,U}$ , one has

$$\begin{aligned}
m(n+1)\mathbb{P}[X_{\ell+1} = Z_{\ell+1}; H' | H_\ell] &\geq 1 - \mathbb{E} \left[ \int_E [f(a|X_\ell^\triangleleft) - M]_+ d\pi(a) \mid H_\ell \right] \\
&\quad - \mathbb{E} \left[ \int_E [f(a|Z_{1:\ell}) - mf(a|X_\ell^\triangleleft)]_+ d\pi(a) \mid H_\ell \right] \\
&\geq 1 - 2\varepsilon/3,
\end{aligned}$$

where the last inequality stands from lemmas 3.2 and 3.3. But

$$\begin{aligned}
\mathbb{P}[X_{\ell+1} = Z_{\ell+1}, H' | H_\ell] &= \mathbb{P}[H' | H_\ell] \mathbb{P}[X_{\ell+1} = Z_{\ell+1} | H' \cap H_\ell] \\
&= \frac{1}{m(n+1)} \mathbb{P}[X_{\ell+1} = Z_{\ell+1} | H_{\ell+1}].
\end{aligned}$$

Hence,

$$\mathbb{P}[Z_{\ell+1} = X_{\ell+1} | H_{\ell+1}] \geq 1 - 2\varepsilon/3.$$

But, since  $H_{\ell+1} = H_\ell \cap H'$  and since  $H'$  is independent of  $\mathcal{F}_\ell^{X,U}$ , the induction hypothesis yields

$$\mathbb{P}[Z_\ell \neq X_{1:\ell} | H_{\ell+1}] = \mathbb{P}[Z_\ell \neq X_{1:\ell} | H_\ell] \leq \varepsilon/3,$$

so

$$\mathbb{P}[Z_{1:\ell+1} \neq X_{1:\ell+1} | H_{\ell+1}] \leq \mathbb{P}[Z_\ell \neq X_{1:\ell} | H_{\ell+1}] + \mathbb{P}[Z_{\ell+1} \neq X_{\ell+1} | H_{\ell+1}] \leq \varepsilon,$$

which ends the proof.  $\square$

## 4 End of the proof of theorem 3

In this section, we assume that the priming condition and condition  $\mathcal{H}'(\eta)$  hold. We show that the parametrization  $(U_n)_{n \in \mathbb{Z}}$  given by proposition 2.5 is generating. By stationarity, it suffices to show that  $X_0$  is a function of  $(U_n)_{n \in \mathbb{Z}}$  only. We proceed by successive approximations.

### 4.1 Approximation until a given time

Choose  $\varepsilon > 0$  and  $\ell \geq 1$  such that  $\sum_{n \geq \ell} \eta_n \leq \varepsilon$ , let  $J = [s, t]$  be an interval of integers such that  $t - s + 1 = \ell$ . Note  $X_J = X_{s:t}$ .

By stationarity of  $X$ , lemma 3.1 provides an event  $H_J$  and a random variable  $Z_J$ , functions of  $U_J$  only, such that

- $\mathbb{P}[H_J] > 0$ ,
- $\mathcal{L}(Z_J) = \mathcal{L}(X_J) = \mathcal{L}(X_{1:\ell})$ ,
- $H_J$  is independent of  $Z_J$ ,
- $\mathbb{P}[X_J \neq Z_J \mid H_J] \leq \varepsilon$ .

Using  $Z_J$  and the parametrization  $(U_n)_{n \geq t+1}$ , we consider the random variables  $(X'_n)_{n \geq s}$  defined by  $X'_J = Z_J$  and for every  $n \geq t+1$ ,

$$X'_n = x_{f'_{n-1}}(U_n) \text{ where } f'_{n-1} = f(\cdot | X'_{s:n-1}).$$

Let us establish some properties of the process  $X'$  thus defined.

**Lemma 4.1** *For every  $n \geq s$ , the law of  $X'_{s:n}$  is the law of  $X_{s:n}$ .*

*Proof.* Choose  $n \geq t+1$ ,  $z \in E^{n-s}$  and  $B \in \mathcal{E}$ . By lemma 2.3 and by independence of  $U_n$  and  $f'_{n-1}$ , the random variable  $X'_n$  admits the density  $f'_{n-1} = f(\cdot | X'_{s:n-1})$  conditionally on  $X'_{s:n-1}$ . Hence

$$\mathbb{P}[X'_n \in B \mid X'_{s:n-1} = z] = \int_B f(x|z) d\pi(x) = \mathbb{P}[X_n \in B \mid X_{s:n-1} = z].$$

Since  $X'_J = Z_J$  has the same law as  $X_J$ , the result follows by induction.  $\square$

**Lemma 4.2** *One has  $\mathbb{P}[X' \neq X \text{ on } [s, +\infty[ \mid H_J] \leq 3\varepsilon$ .*

*Proof.* Let  $n \geq t + 1$ . Since  $X_n = x_{f_{n-1}}(U_n)$  and  $X'_n = x_{f'_{n-1}}(U_n)$ , proposition 2.3 and the independence of  $U_n$  and  $\mathcal{F}_{n-1}^{X,U}$  yield,

$$\mathbb{P}[X'_n \neq X_n \mid \mathcal{F}_{n-1}^{X,U}] \leq 2\|f'_{n-1} - f_{n-1}\|.$$

Let

$$p_n = \mathbb{P}[X'_n \neq X_n ; X'_{s:n-1} = X_{s:n-1} ; H_J].$$

Since  $\{X'_{s:n-1} = X_{s:n-1} ; H_J\} \in \mathcal{F}_{n-1}^{X,U}$ , one gets

$$\begin{aligned} p_n &\leq \mathbb{E}\left[2\|f_{n-1} - f'_{n-1}\| \mathbf{1}_{\{X'_{s:n-1} = X_{s:n-1}\}} \mathbf{1}_{H_J}\right] \\ &= 2 \mathbb{E}\left[\|f(\cdot | X_{s-1}^\triangleleft X'_{s:n-1}) - f(\cdot | X'_{s:n-1})\| \mathbf{1}_{\{X'_{s:n-1} = X_{s:n-1}\}} \mathbf{1}_{H_J}\right] \\ &\leq 2 \mathbb{E}\left[\|f(\cdot | X_{s-1}^\triangleleft X'_{s:n-1}) - f(\cdot | X'_{s:n-1})\| \mathbf{1}_{H_J}\right]. \end{aligned}$$

But  $X_{s-1}^\triangleleft$ ,  $Z_J$ ,  $H_J$  and  $U_{t+1:n-1}$  are independent hence  $X_{s-1}^\triangleleft$ ,  $X'_{s:n-1}$  and  $H_J$  are independent since  $X'_{s:n-1}$  is a function of  $Z_J$  and  $U_{t+1:n-1}$  only. Thus,

$$p_n \leq 2 \mathbb{E}\left[\|f(\cdot | X_{s-1}^\triangleleft X'_{s:n-1}) - f(\cdot | X'_{s:n-1})\|\right] \mathbb{P}[H_J] = 2\eta_{n-s}\mathbb{P}[H_J].$$

Hence,

$$\mathbb{P}[X'_n \neq X_n ; X'_{s:n-1} = X_{s:n-1} \mid H_J] \leq 2\eta_{n-s},$$

therefore,

$$\mathbb{P}[X'_{s:n} \neq X_{s:n} \mid H_J] \leq \mathbb{P}[X'_{s:n-1} \neq X_{s:n-1} \mid H_J] + 2\eta_{n-s}.$$

By induction, one gets for all  $n \geq t + 1$

$$\mathbb{P}[X'_{s:n} \neq X_{s:n} \mid H_J] \leq \mathbb{P}[X'_J \neq X_J \mid H_J] + 2 \sum_{m=\ell}^{n-s} \eta_m.$$

Since  $X'_J = Z_J$  and  $\mathbb{P}[X_J \neq Z_J \mid H_J] \leq \varepsilon$ , this yields

$$\mathbb{P}[X' \neq X \text{ on } [s, +\infty[ \mid H_J] \leq \varepsilon + 2 \sum_{m=\ell}^{\infty} \eta_m \leq 3\varepsilon,$$

which ends the proof.  $\square$

## 4.2 Successive approximations

Our next step in the proof of theorem 3 is to approach the random variable  $X_0$  by measurable functions of the parametrization  $(U_n)_{n \in \mathbb{Z}}$  of proposition 2.5. To this aim, we group the innovations by intervals of times. For every  $m \geq 1$  one chooses an integer  $L_m$  such that

$$\sum_{n \geq L_m} \eta_n \leq 1/m.$$

Lemma 3.1 (the priming lemma) applied to  $\ell = L_m$  and  $\varepsilon = 1/m$  provides an event  $H_{L_m}$  of positive probability  $\alpha_m$  and a random variable  $Z_{L_m}$  such that

$$\mathbb{P}[X_{1:L_m} = Z_{L_m} \mid H_{L_m}] \geq 1 - 1/m.$$

Choose an integer  $M_m \geq 1/\alpha_m$ . Split  $\mathbb{Z}_-$  into  $M_1$  intervals of length  $L_1$ ,  $M_2$  intervals of length  $L_2$ , and so on. More precisely set, for every  $n \geq 1$ ,  $\ell_n = L_{m(n)}$ ,  $\varepsilon_n = 1/m(n)$  and  $\alpha_n = \alpha_{m(n)}$  where  $m(n)$  is the only integer such that

$$M_1 + \dots + M_{m(n)-1} < n \leq M_1 + \dots + M_{m(n)}.$$

Therefore, for every  $k \geq 0$ ,

$$\sum_{n \geq \ell_k} \eta_n \leq \varepsilon_k.$$

For every  $k \geq 0$ , set

$$t_k = - \sum_{1 \leq n \leq k} \ell_n, J_k = [t_k, t_k + \ell_k - 1] = [t_k, t_{k-1} - 1] \text{ and } X_{J_k} = X_{t_k : t_{k-1} - 1}.$$

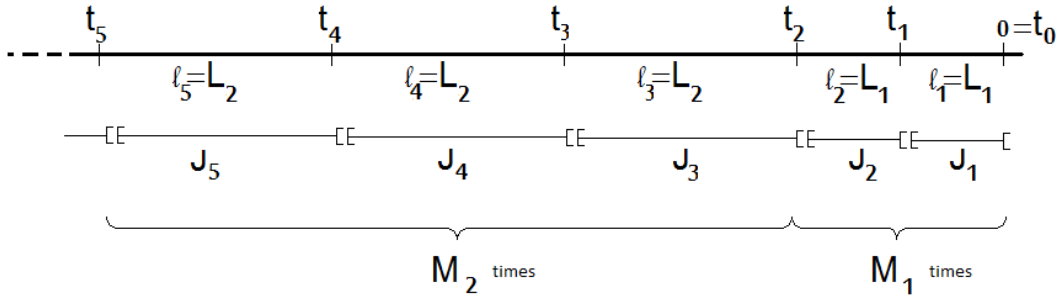


Figure 1: Splitting  $\mathbb{Z}_-$  into intervals of times.

Lemma 3.1 applied to  $(\varepsilon_k)_{k \geq 1}$  and  $(U_{J_k})_{k \geq 1}$  provides events  $(H_{J_k})_{k \geq 1}$  and random variables  $(Z_{J_k})_{k \geq 1}$ . For every  $k \geq 0$ , let us use the construction of section 4.1: set  $X_{J_k}^k = Z_{J_k}$ , then for every  $n \geq t_k + \ell_k = t_{k-1}$

$$X_n^k = x_{f_{n-1}^k}^k(U_n) \text{ where } f_{n-1}^k = f(\cdot | X_{t_k:n-1}^k).$$

Therefore, lemma 4.2 yields the inequality

$$\mathbb{P}[X_0 \neq X_0^k \mid H_{J_k}] \leq \mathbb{P}[X_{t_k:0} \neq X_{t_k:0}^k \mid H_{J_k}] \leq 3\varepsilon_k,$$

which shows that

$$\mathbb{P}[X_0^k \neq X_0 \mid H_{J_k}] \rightarrow 0 \text{ when } k \rightarrow +\infty.$$

Moreover, the events  $H_{J_k}$  are independent as functions of random variables  $U_j$  for disjoint sets of indices  $j$  and

$$\sum_{k \geq 1} \mathbb{P}[H_{J_k}] = \sum_{n \geq 1} \alpha_{\ell_{m(n)}} = \sum_{m=1}^{+\infty} M_m \alpha_m = +\infty$$

since  $M_m \alpha_m \geq 1$  by choice of  $M_m$ .

Lemma 4.3, stated below, provides a deterministic increasing function  $\theta$  such that

$$\sum_{k \geq 1} \mathbb{P}[X_0^{\theta(k)} \neq X_0 ; H_{J_{\theta(k)}}] < +\infty$$

$$\text{and } \sum_{k \geq 1} \mathbb{P}[H_{J_{\theta(k)}}] = +\infty.$$

Using Borel-Cantelli's lemma, one deduces that

- $\{X_0^{\theta(k)} \neq X_0\} \cap H_{J_{\theta(k)}}$  occurs only for finitely many  $k$  only, a.s.
- $H_{J_{\theta(k)}}$  occurs for infinitely many  $k$  a.s.

Thus, for every  $B \in \mathcal{E}$ ,

$$\{X_0 \in B\} = \limsup_{k \rightarrow \infty} H_{J_{\theta(k)}} \cap \{X_0^{\theta(k)} \in B\},$$

hence  $\{x_0 \in B\} \in \mathcal{F}_0^U$ . By stationarity of the process  $(X, U)$ , one gets the inclusion of the filtration  $\mathcal{F}^X$  into the filtration  $\mathcal{F}^U$ , which ends the proof.  $\square$

**Lemma 4.3** *Let  $(a_n)_{n \geq 0}$  and  $(b_n)_{n \geq 0}$  denote two bounded sequences of nonnegative real numbers such that the series  $\sum_n b_n$  diverges and such that  $a_n \ll b_n$ . Then there exists an increasing function  $\theta : \mathbb{N} \rightarrow \mathbb{N}$  such that the series  $\sum_n a_{\theta(n)}$  converges and the series  $\sum_n b_{\theta(n)}$  diverges.*

A proof of this lemma can be found in [4].

## 5 proof of proposition 1.2

Assume that  $E$  is countable. Let  $\varepsilon = (\varepsilon_a)_{a \in E}$  be an i.i.d. family of exponential random variables with parameter 1.

Let us show that, for any probability  $p$  on  $E$ , there exists almost surely one and only one  $a \in E$  such that

$$\frac{\varepsilon_a}{p(a)} = \inf_{b \in E} \frac{\varepsilon_b}{p(b)} \quad (\star),$$

with the convention that  $\varepsilon_b/p(b) = +\infty$  if  $p(b) = 0$ . For every positive real number  $r$ ,

$$\sum_{b \in E} \mathbb{P}[\varepsilon_b/p(b) \leq r] = \sum_{b \in E} (1 - e^{-p(b)r}) \leq \sum_{b \in E} p(b)r = r < +\infty,$$

hence Borel-Cantelli's lemma ensures that the event  $\{\varepsilon_b/p(b) \leq r\}$  occurs only for finitely many  $b \in E$ . Thus the infimum in  $(\star)$  is achieved at some  $a \in E$ . The uniqueness follows from the equalities  $\mathbb{P}[\varepsilon_a/p(a) = \varepsilon_b/p(b) < +\infty] = 0$  for every  $a \neq b$ , since  $\varepsilon_a$  and  $\varepsilon_b$  are independent random variables with diffuse laws.

Define  $g(\varepsilon, p)$  (almost surely) as the only index  $a$  such  $(\star)$  holds. Let us show that, for every  $p, q \in \mathcal{P}(E)$ ,

$$\mathbb{P}[g(\varepsilon, p) \neq g(\varepsilon, q)] \leq 2 \frac{\|p - q\|}{1 + \|p - q\|}.$$

For every  $a, b \in E$ , set  $C_a = \{g(\varepsilon, p) = g(\varepsilon, q) = a\}$  and

$$\lambda_{b/a} = \max\left(\frac{p(b)}{p(a)}, \frac{q(b)}{q(a)}\right).$$

Fix  $a \in E$ . Then up to negligible events,

$$\{g(\varepsilon, p) = a\} = \bigcap_{b \neq a} \left\{ \frac{\varepsilon_a}{p(a)} \leq \frac{\varepsilon_b}{p(b)} \right\} = \bigcap_{b \neq a} \left\{ \varepsilon_b \geq \frac{p(b)}{p(a)} \varepsilon_a \right\},$$

and the same holds for  $q$ . Thus,

$$C_a = \bigcap_{b \neq a} \left\{ \varepsilon_b \geq \lambda_{b/a} \varepsilon_a \right\},$$

Conditioning on  $\varepsilon_a$  and using the fact that the random variables  $(\varepsilon_b)$  are i.i.d. and exponentially distributed, one gets

$$\mathbb{P}[C_a \mid \varepsilon_a] = \prod_{b \neq a} \exp(-\lambda_{b/a} \varepsilon_a),$$

hence

$$\mathbb{P}(C_a) = \mathbb{E} \left[ \exp \left( - \sum_{b \neq a} \lambda_{b/a} \varepsilon_a \right) \right] = \left( 1 + \sum_{b \neq a} \lambda_{b/a} \right)^{-1} = \left( \sum_b \lambda_{b/a} \right)^{-1}. \quad (3)$$

But  $\lambda_{b/a} \leq \max(p(b), q(b)) / \min(p(a), q(a))$ , hence

$$\mathbb{P}(C_a) \geq \frac{\min(p(a), q(a))}{\sum_b \max(p(b), q(b))} = \frac{\min(p(a), q(a))}{1 + \|p - q\|}.$$

Therefore

$$\mathbb{P}[g(\varepsilon, p) = g(\varepsilon, q)] = \sum_{a \in E} \mathbb{P}(C_a) = \frac{1 - \|p - q\|}{1 + \|p - q\|}.$$

Last, note that if  $p = q$ , then for each  $a \in E$ , equation (3) becomes

$$\mathbb{P}[g(\varepsilon, p) = a] = \mathbb{P}(C_a) = \left( \sum_b \frac{p(b)}{p(a)} \right)^{-1} = p(a),$$

which shows that the law of  $g(\varepsilon, p)$  is  $p$ . The proof is complete.  $\square$

## References

- [1] X. Bressaud, A. Maass, S. Martinez, and J. San Martin. *Stationary processes whose filtrations are standard*. Annals of Probability **34**, no 4, 1589–1600 (2006).
- [2] G. Ceillier, *Filtrations à temps discret*. Université Joseph Fourier, Grenoble, France (2010).
- [3] G. Ceillier, *Sufficient conditions of standardness for filtrations of finite stationary process*. Annals of Probability **40**, no 5, 1980–2007 (2012).
- [4] G. Ceillier, *The filtration of the split-words process*. Probability Theory and Related Fields **153**, no 1-2, 269–292 (2012).
- [5] F. Comets, R. Fernandez, and P.A. Ferrari. *Processes with long memory: regenerative construction and perfect simulation*. Annals of Applied Probability **12**, no 3, 921–943 (2002).
- [6] M. Émery and W. Schachermayer. *On Vershik’s standardness criterion and Tsirelson’s notion of cosiness*. Séminaire de Probabilités, XXXV, LNM **1755**, 265–305 (2001).

- [7] D. Heicklen and C. Hoffman.  $[T, T^{-1}]$  is not standard. *Ergodic Theory and Dynamical Systems* **18**, no 4, 875–878 (1998).
- [8] S. Laurent, *Filtrations à temps discret négatif*. PhD thesis, Université Louis Pasteur, Institut de Recherche en Mathématique Avancée, Strasbourg (2004).
- [9] W. Schachermayer. *On certain probabilities equivalent to Wiener measure, d'après Dubins, Feldman, Smorodinsky and Tsirelson*. Séminaire de Probabilités, XXXIII, LNM **1709**, 221–239 (1999). See also the addendum in Séminaire de Probabilités, XXXVI, LNM **1801**, 493–497 (2002).
- [10] M. Smorodinsky, *Processes with no standard extension*. *Israel Journal of Mathematics*, **107**, 327–331 (1998).
- [11] A. Vershik, *Theory of decreasing sequences of measurable partitions*. *Algebra i Analiz*, **6**, no 4, 1–68 (1994). English Translation: *St. Petersburg Mathematical Journal*, **6**, no 4, 705–761 (1995).